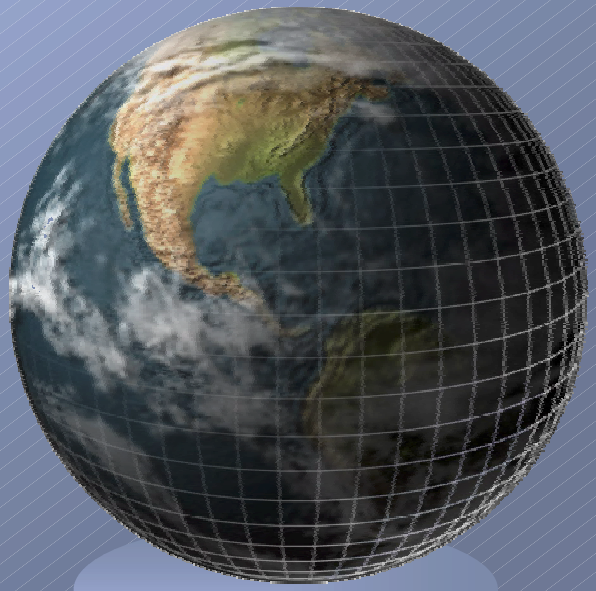
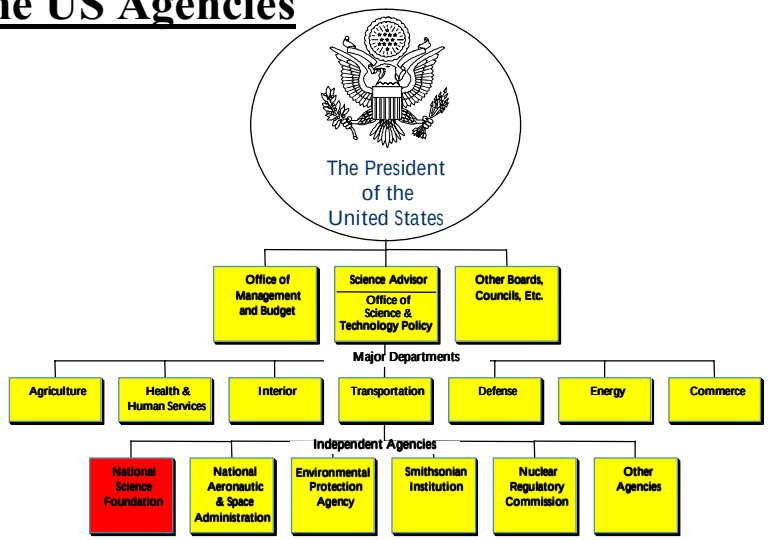
CyberInfrastructure (CI): Whence?



Thomas J. Greene, Ph.D.
Sr. Program Dir. for CISE-ANIR,
National Science Foundation

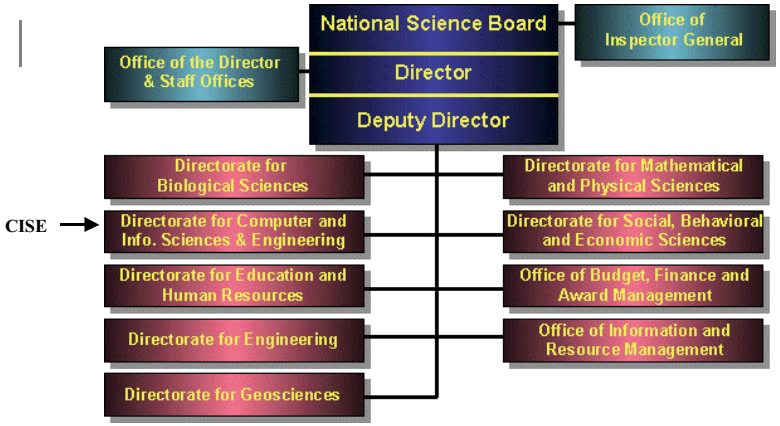
The US Agencies



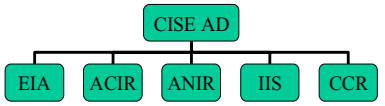
What Is NSF?

- **An Independent Agency of the Federal Government, established in 1950,**
  - **to promote and advance scientific progress in the United States**
  - **by sponsoring scientific research and**
  - **by supporting selected activities in science and engineering.**
- **NSF Manages the US Science & engineering research investment portfolio**
- **NSF Does not conduct research itself**

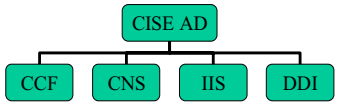
National Science Foundation & CISE



CISE Division Structure

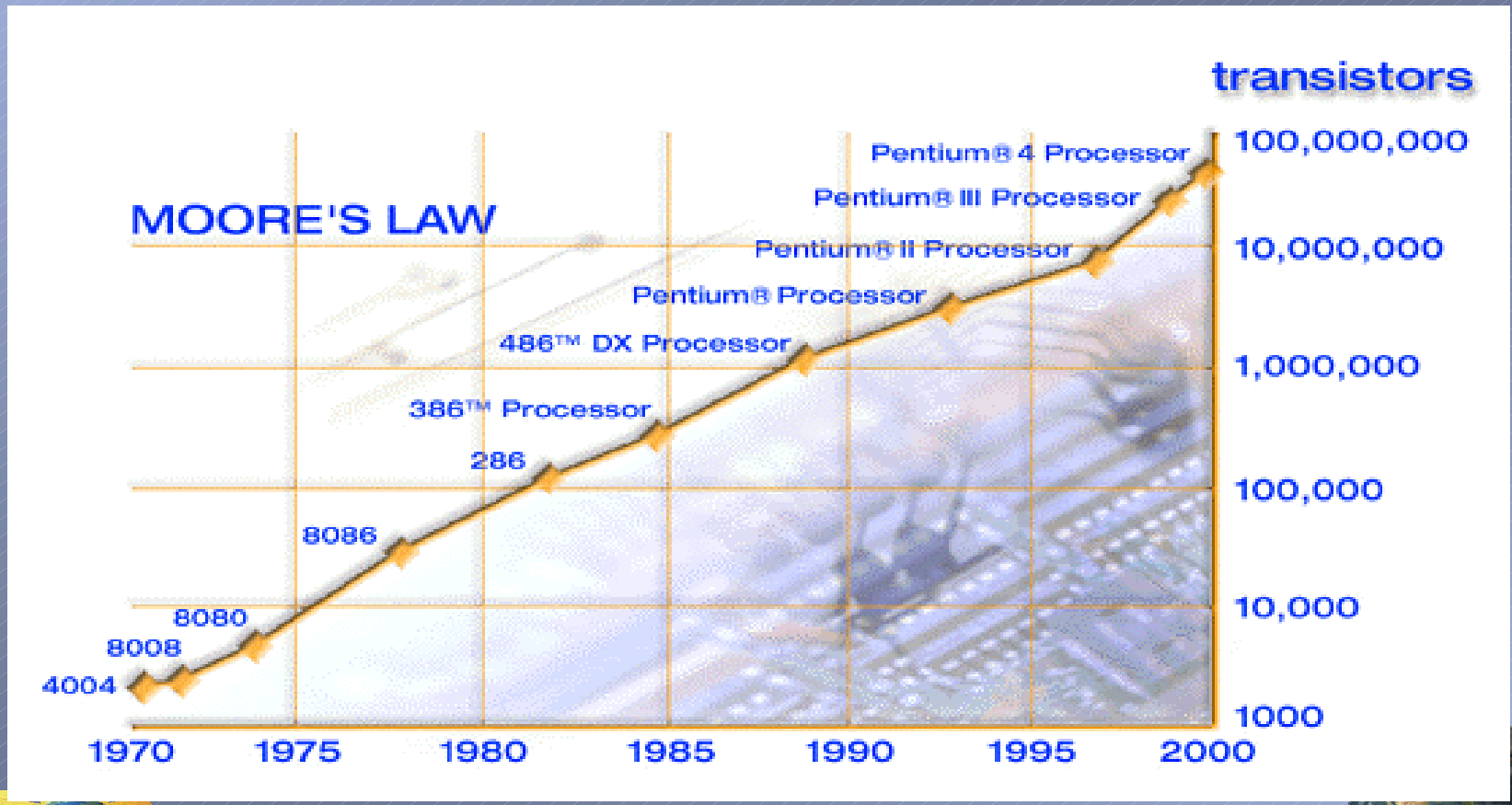


Reorganization – Effective 1/OCT/03



Proposed CISE Organization

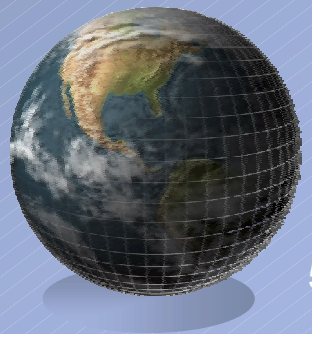
The drivers – *Faster cheaper power & networks*



Moore on cost & Metcalf on value

Outline

1. The state of NSF CISE vision of CI
2. What strategy with which to proceed



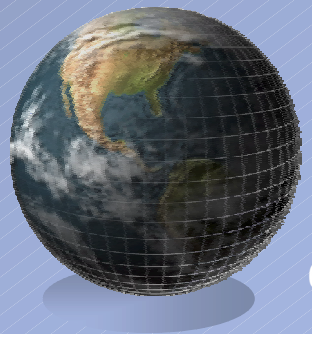
Great Big Idea: Cyberinfrastructure (CI)

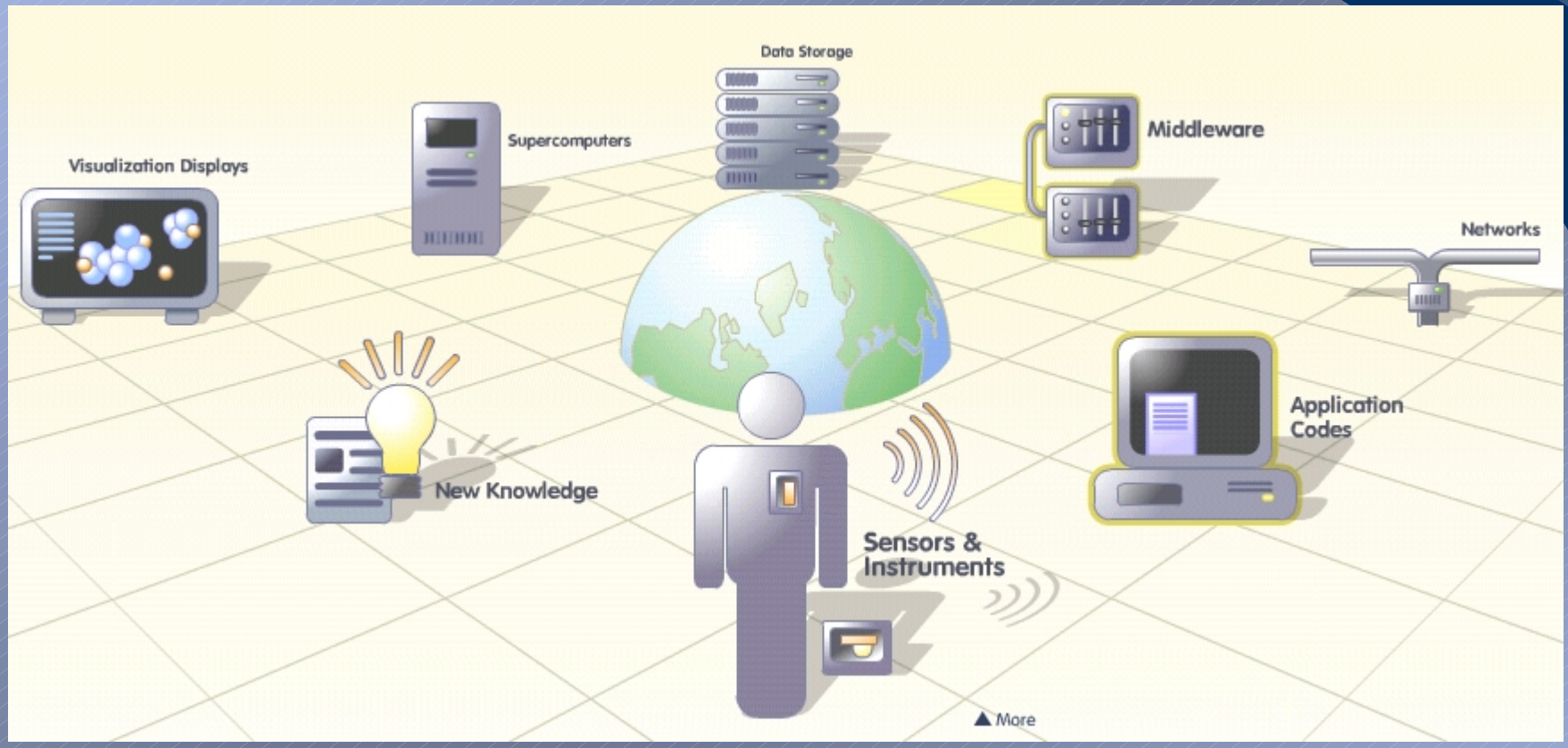
A "Scaling up" of some older ideas

The old components are

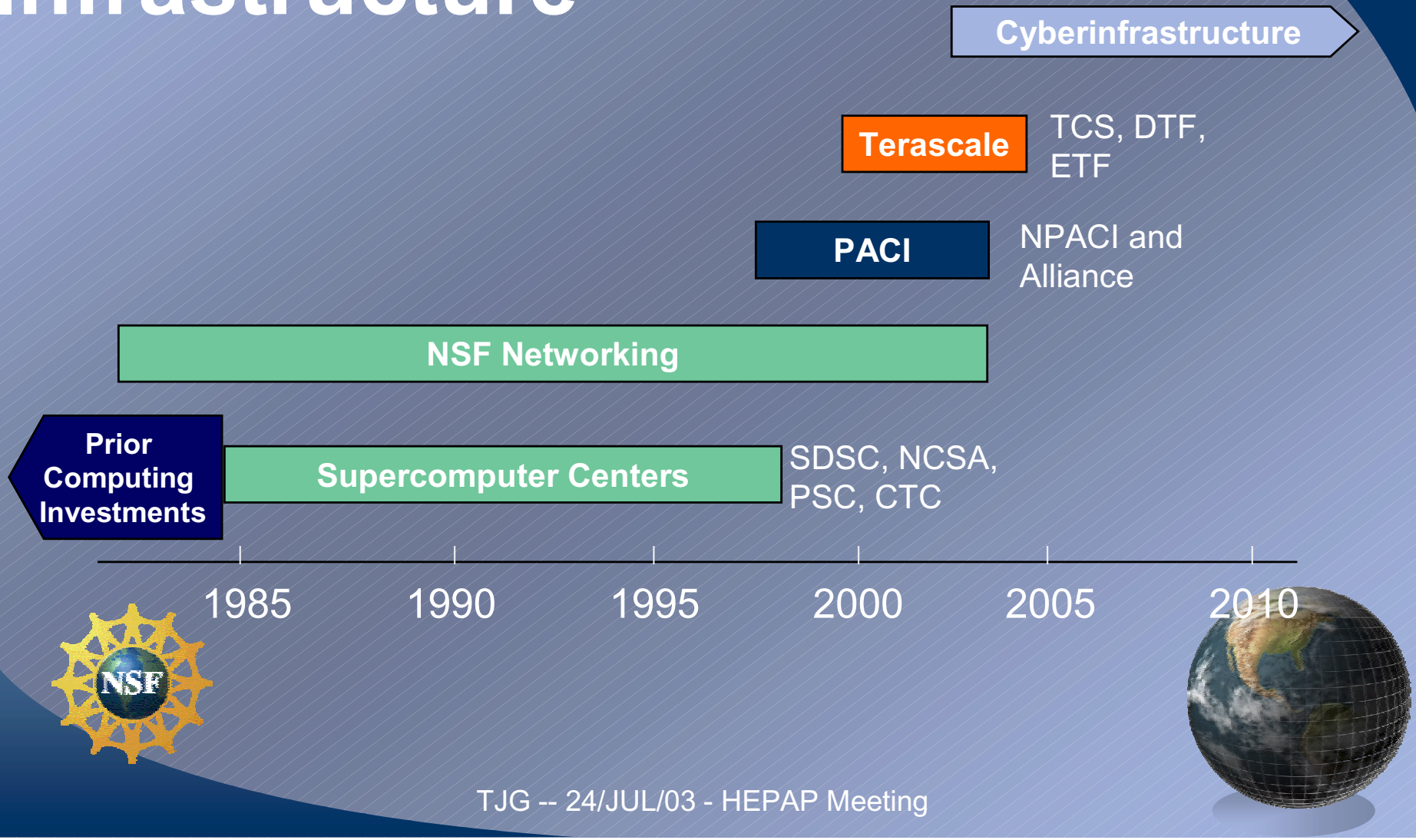
- INPUT/OUTPUT devices
- PROCESSORS
- MEMORY
- CONNECTIONS
- DATA
- PROGRAMS

Today- very very powerful and GLOBAL is CI



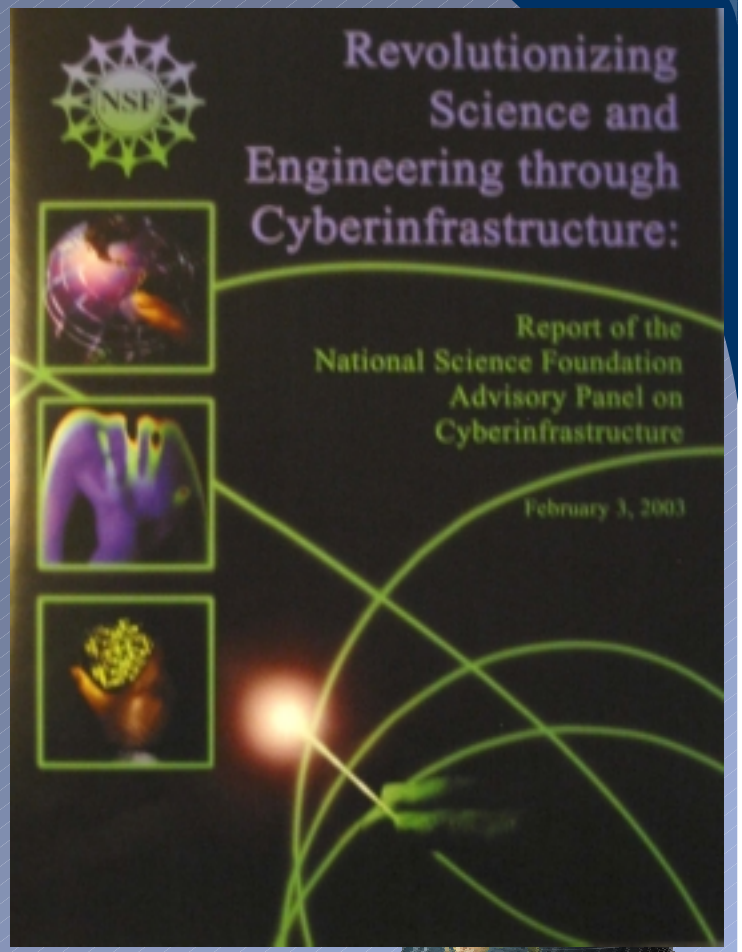

Data Storage
Supercomputers
Visualization Displays
Middleware
Networks
Application Codes
New Knowledge
Sensors & Instruments
More

Evolution of the NSF programs in Computational Infrastructure



Setting the Stage

Daniel E. Atkins, Chair, University of Michigan
Kelvin K. Droegemeier, University of Oklahoma
Stuart I. Feldman, IBM
Hector Garcia-Molina, Stanford University
Michael L. Klein, University of Pennsylvania
David G. Messerschmitt, University of California at Berkeley
Paul Messina, California Institute of Technology
Jeremiah P. Ostriker, Princeton University
Margaret H. Wright,New York University

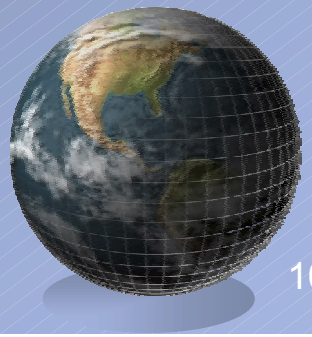


http://www.communitytechnology.org/nsf_ci_report/

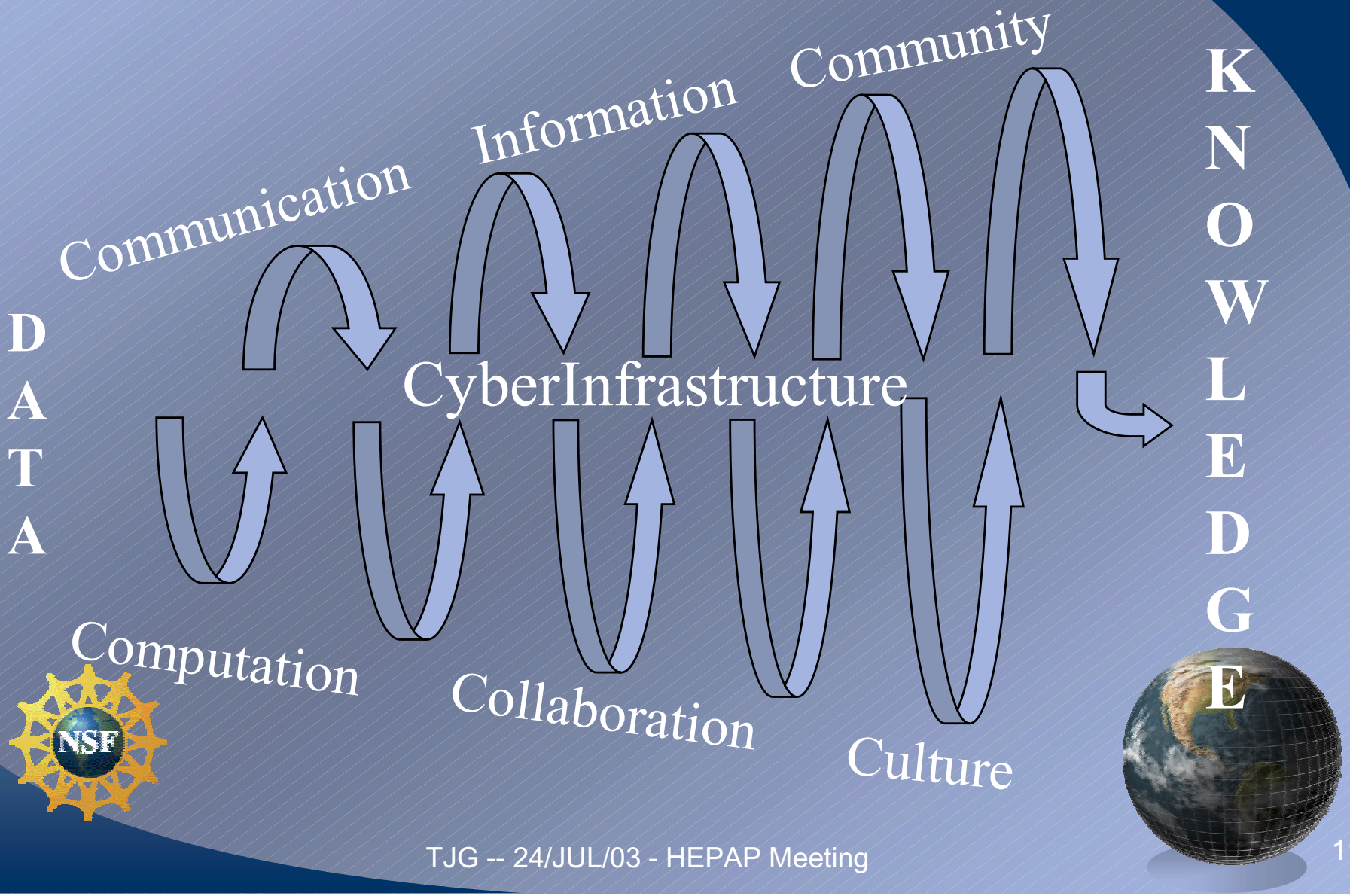
The CI Promise

- Ubiquitous, digital knowledge environments that are both interactive and functionally complete……

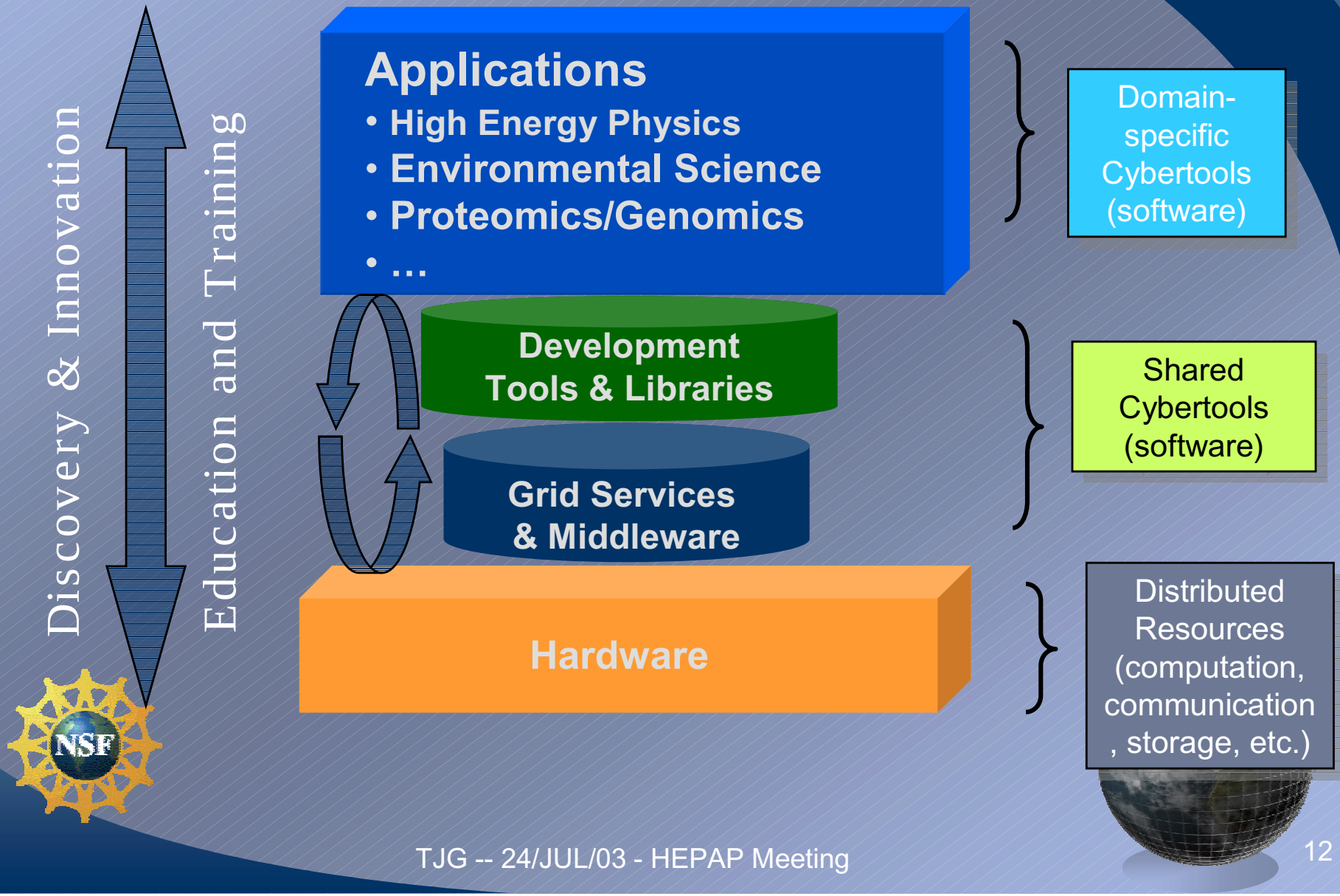
- Revolutionize the processes of discovery, learning and innovation across the science and engineering frontier.



On The Path to Knowledge

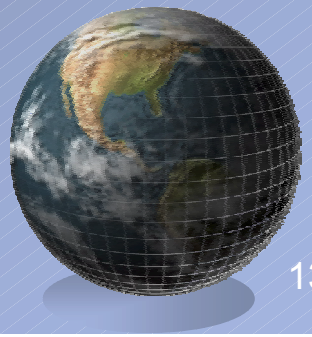


Integrated CI System: meeting the needs of a community of communities



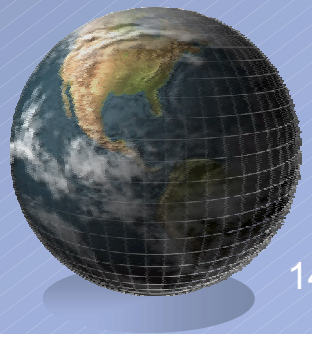
Characteristics of CyberInfrastructure

- Community-Driven
- Distributed Collaboration
- Virtual Organization
- Multidisciplinary in scale and scope
- International in scale and scope
- Interoperable
- Supporting Data- and Compute-Intensive Applications
- High end to desktop
- Distributed
- Heterogeneous
- Complex
- Reusable



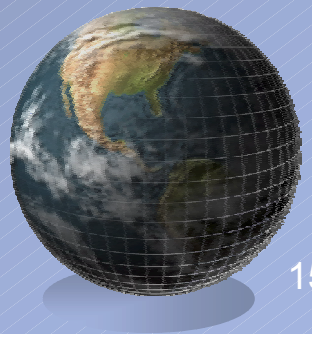
NSF-CISE Plan of Action

- Focused, cross-cutting attention on CyberInfrastructure – not business as usual
- Internal NSF planning now underway - active discussion on specific CyberInfrastructure issues
- Community building - broad consultations with scientific communities will intensify in coming months
- Summer 2003 workshops and town hall meetings – management models
- NSF FY05 budget planning for CyberInfrastructure beginning shortly



2. How can we Proceed?

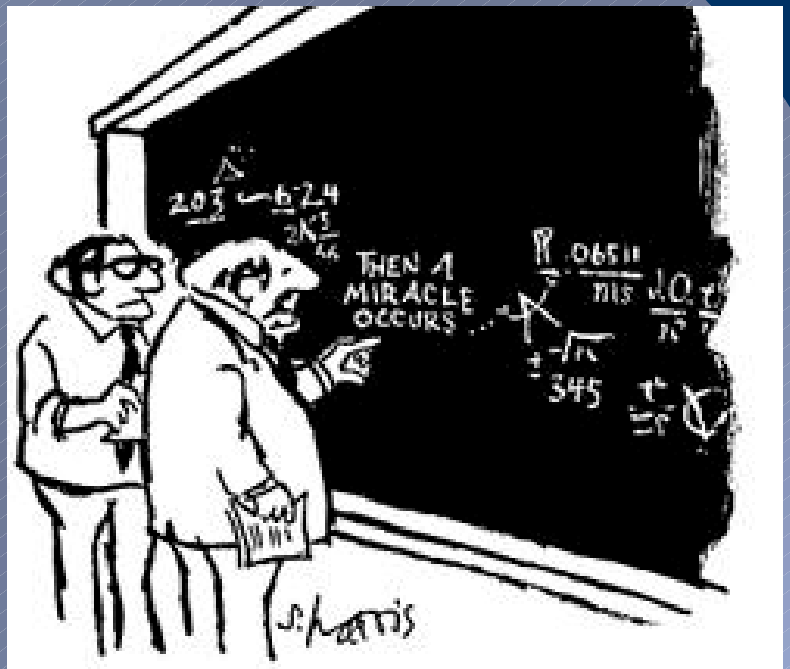
- Big SCIENCE meets Big IT.
- Overlap the native interests of the natural duality of "IT Builders communities" and "IT Users communities" to rapidly advance CI for both.
- The CI Community of Communities



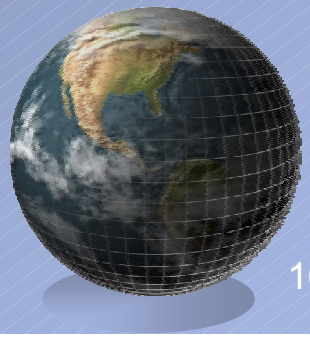
Cyberinfrastructure is a Large COMPLEX system – (with Synergy?)

- *top down*
- *bottom up*

- *or middle out*



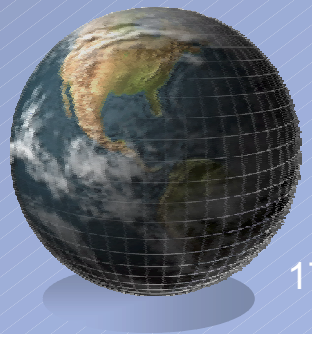
"I think you should be more explicit here in step two."



INTEROPERABILITY

- Biology, Physics, Chemistry, Geological Science, Environmental Science, and others
- Interoperability of Tools for multiple domains of science should be kept in view—
- E-science involves data, measurement etc. so whenever possible play “open” -- then specialize

- Collaboration (NAPs?)

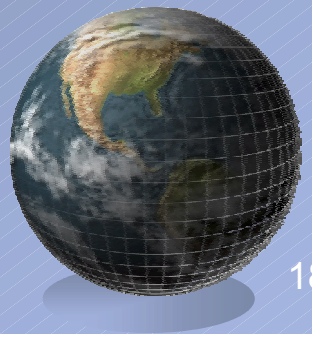


A Continuing Discussion and evolving Collaboration of –
Computer & High Energy Physics
People with the ***Idea*** of ***Tools*** to share

Communities involved in building GLOBAL INTERAGENCY COLLABORATION are:

DOE: OASCR; HEP NSF: CISE; EPP

CERN; EU *More …*

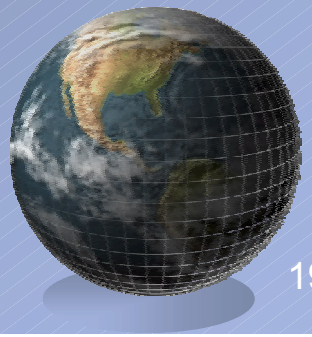


Why would they play together?

- Could be:
  - (1) Altruism (help the other guy)
  - (2) Self- Interest (I need IT)

Choose 2 because

- Busy communities need success in their own agenda … so they more easily respond to Self-interest (collaboration to achieve Leveraged outcomes of their agendas)



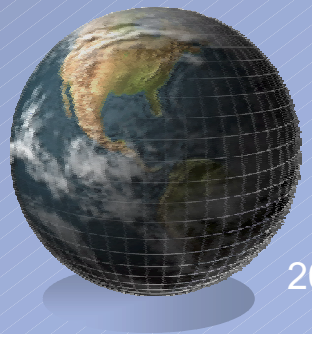
A snapshot of a day some time in the future

Some problem solver is thinking about a problem and goes to work ---

- *Gathers all known data from around the globe*
- *Transforms data in several ways using remote engines*
- *Looks at both initial and transformed data in several ways*
- *Dreams a model solution and runs a simulation predicting an new outcome*
- *Makes a new measurement confirming the model*
- *Consults with colleagues*
- *Claims the solution as a breakthrough and broadcasts it*

This person does this while touching any spot on the globe in real time

WHEN?

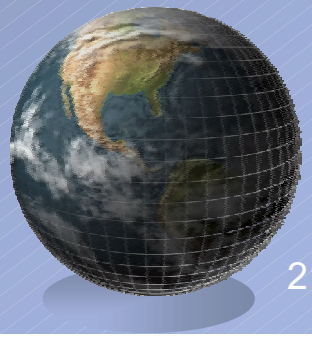


Let's Contract This Expansive Vision to a Smaller Space

Use a strategy of implementing this as much as possible with existing known IT components

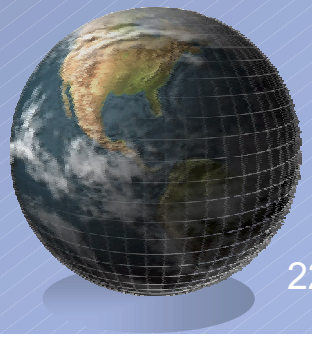
- **Build new pieces as needed**
- **Make some "standards choices" to enable the whole system and advance interoperability**
- **Find "risk takers" committed to clear goals who are motivated to collaborate (when the leveraged gain for their goal is visible,)**

Define the visible common goal, assemble the goal seekers and overlap the resources



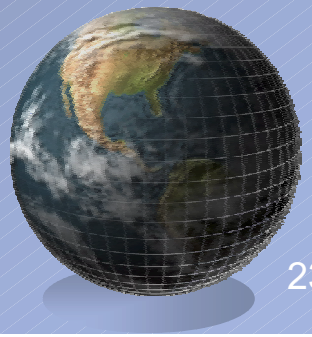
Bad News L

- **Crossing cultures to enable collaborations is not a well understood process.**
- **New tools for this are needed in computers and communications.**
- **A fast agreement to enabling standards will require people choosing to agree to further the big agenda – "big thinkers".**



Good News - *J* a Win-win Situation Exists

- **Essential changes in information access are happening now, (not just a little faster, bigger but - much faster, bigger, wider!).**
- **Solving very hard problems faster and better by the collaborations "idea" by some frontiersmen is already agreed to .**
- **Committed problem solvers will "climb mountains " to solve their problem agenda and will even work with people outside their culture – *sometimes.***



Some Principles for a Global Cyber Infrastructure and for E-science Rapid Prototype

Five principles for Cyber infrastructure and LHC communities :

0] The cost and complexity of 21st Century Science requires the creation of advanced and coherent global Infostructure (information infrastructure).

1] The construction of a coherent Global Infostructure for Science requires definition and drivers from Global Applications (that will also communicate with each other)

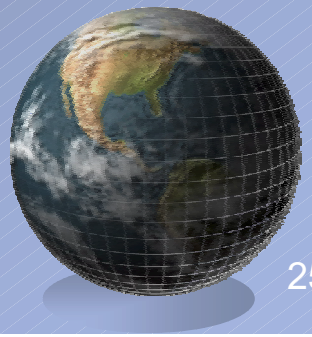
2] Further, forefront Information Technology must be incorporated into this Global Infostructure for the Applications to reach their full potential for changing the way science is done

3] LHC is a near term Global Application requiring advanced and un-invented Infostructure and is ahead in planning compared to many others

4] U.S. agencies must work together for effective U.S. participation on Global scale infostructure and the successful execution of the LHC program in a 4 way agency partnership, with international cooperation in view

Strategy for CI building

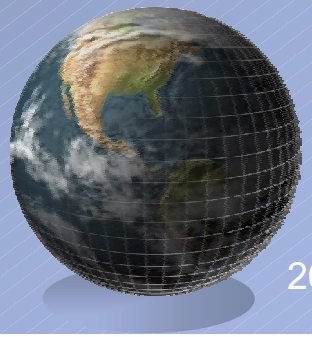
- Pick some existing global big science projects
- Touch them with projects building cyber infrastructure components
- Agree to force “de facto standards” for mutual benefit across cultural divides



CI Building Tactics

- Use common agendas, not common check books.(*i.e. Keep money inside the native projects/agencies as much as possible*.)

- Full credit to all
- Give Endorsements across cultures whenever useful.

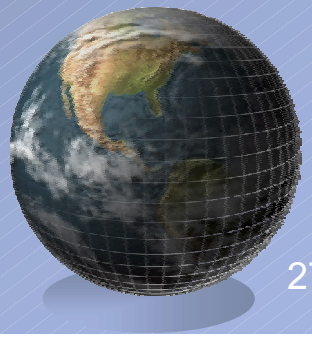


CI Is a Large system so build it…

- Bottom up --- micro
- Top down --- macro
- Middle out – engineering choices for standard interfaces (Standards solutions etc.)

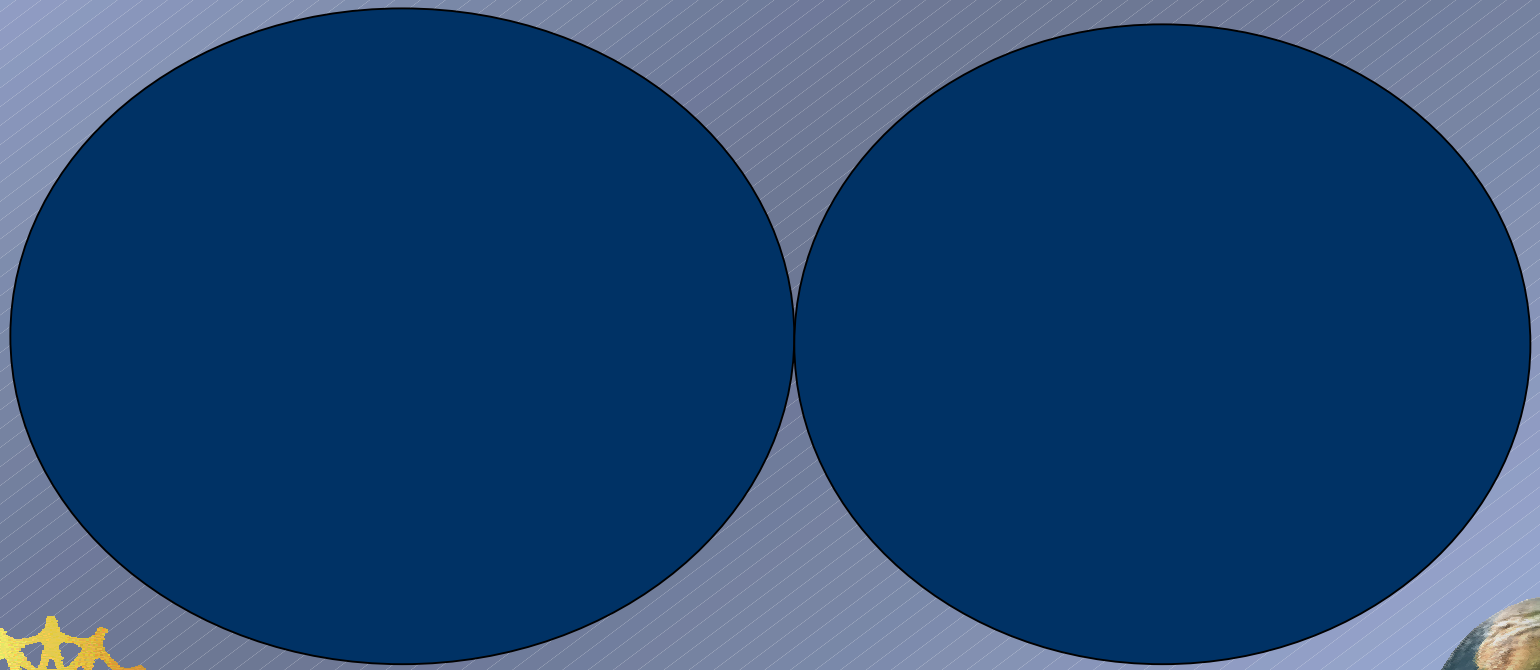
--

- USE Cyber infrastructure rapid proto demos
- And keep them alive after the demo …



Leverage at the Interface

- SCIENCE APPLICATIONS
- COMPUTER TOOLS



Deus ex machina?

"You can't just punch in 'let there be light' without writing the code underlying the user interface functions."

A global vision of Ubiquitous information at Light-speed
– Cyberinfrastucture (Grids, E-science) =